

Folha de Pagamento - SPOBRAS 2019

```
library(tidyverse)
library(na.tools)
library(readxl)
library(stringr)
library(lubridate)
library(pdftools)
library(data.table)
```

```
spobras<-fread("http://dados.prefeitura.sp.gov.br/dataset/3edd72cc-2f79-4c95-88e7-55744881b0a6/resource/898bfaa8-6583-4ea4-9a20-84b2b5bc6537/download/spobras.csv",sep = "auto",sep2 = "auto",integer64 = "numeric",encoding = "Latin-1")
```

Folha de pagamento SPOBRAS - 2019

```
spobras<-as_tibble(spobras)
spobras
```

EMPRESA	MES	ANO	NOME	
<chr>	<chr>	<int>	<chr>	
SPObras	Janeiro	2019	WALTER MARTINS DE SOUZA	
SPObras	Janeiro	2019	LUIZ CARLOS DE SOUZA BARBOZA	
SPObras	Janeiro	2019	VALDELI DA SILVA	
SPObras	Janeiro	2019	MARCIA APARECIDA TEIXEIRA DE CAMPOS CARVALHO	
SPObras	Janeiro	2019	MARCOS SILVA ROCHA	
SPObras	Janeiro	2019	JOAO PAULO RESPLANDES DE OLIVEIRA	
SPObras	Janeiro	2019	ELIANA MARIANO FRANCO	
SPObras	Janeiro	2019	SILVANA ISGROI CARVALHO	
SPObras	Janeiro	2019	MARCIA HELENA NUCCI EIRAS GARCIA	
SPObras	Janeiro	2019	ALFREDO LENZ	
1-10 of 2,362 rows 1-4 of 15 columns				Previous 1 2 3 4 5 6 ... 237 Next

CONTAGEM DE VALORES FALTANTES EM NOSSO DATASET.

```
sum(is.na(spobras))
```

```
## [1] 0
```

EXCLUINDO A VARIÁVEL LIMINAR

```
spobras<-spobras[, -15]
```

QUANTIDADE DE SERVIDORES EM NOSSO DATASET SPOBRAS

```
spobras %>% distinct(NOME)# RETORNA 241 NOMES
```

```
NOME
<chr>
WALTER MARTINS DE SOUZA
LUIZ CARLOS DE SOUZA BARBOZA
VALDELI DA SILVA
MARCIA APARECIDA TEIXEIRA DE CAMPOS CARVALHO
MARCOS SILVA ROCHA
JOAO PAULO RESPLANDES DE OLIVEIRA
ELIANA MARIANO FRANCO
```

NOME
<chr>
SILVANA ISGROI CARVALHO
MARCIA HELENA NUCCI EIRAS GARCIA
ALFREDO LENZ
1-10 of 241 rows
Previous 1 2 3 4 5 6 ... 25 Next

QUANTIDADE DE CARGOS DIFERENTES NO DATASET

spobras %>% distinct(CARGO) # RETORNA 44 CARGOS DIFERENTES
--

CARGO
<chr>
ANALISTA DE PROJETOS, OBRAS E SERVIÇOS-ENGENHEIRO
AGENTE DE PROJETOS, OBRAS E SERVIÇOS
AGENTE DE ADMINISTRAÇÃO
ANALISTA DE ADMINISTRAÇÃO
SECRETARIA
ASSISTENTE DE PROJETOS OBRAS E SERVIÇOS
ANALISTA DE AGRIMENSURA E CONSTRUÇÃO CIVIL
AUXILIAR DE ADMINISTRAÇÃO
ANALISTA DE PROJETOS OBRAS E SERVIÇOS - ARQUITETO
ADVOGADO
1-10 of 44 rows
Previous 1 2 3 4 5 Next

IDENTIFICANDO A QUANTIDADE DE VINCULOS NO DATASET

spobras %>% group_by(`DETALHE VÍNCULO`) %>% count(`DETALHE VÍNCULO`)
--

DETALHE VÍNCULO	n
<chr>	<int>
Concursado NÃO ocupante de cargo em comissão	1371
Concursado ocupante de cargo em comissão	43
Não concursado NÃO ocupante de cargo em comissão	444
Não concursado ocupante de cargo em comissão	504
4 rows	

TRANSFORMANDO VARIÁVEIS VENCIMENTOS, ENCARGOS E BENEFÍCIOS EM VARIÁVEIS NUMÉRICAS

<pre>### VARIÁVEIS: VENCIMENTOS, ENCARGOS E BENEFÍCIOS ESTÃO COMO strings.. É PRECISO CONVERTE-LAS A NUMEROS ## (substituir a vírgula por ponto) spobras\$VENCIMENTOS<-str_replace_all(spobras\$VENCIMENTOS,pattern = "[,]",replacement = ".") spobras\$ENCARGOS<-str_replace_all(spobras\$ENCARGOS,pattern = "[,]",replacement = ".") spobras\$BENEFÍCIOS<- str_replace_all(spobras\$BENEFÍCIOS,pattern = "[,]",replacement = ".") ### agora transforma-las em numeros com casas decimais spobras\$VENCIMENTOS<-as.double(spobras\$VENCIMENTOS) spobras\$ENCARGOS<-as.double(spobras\$ENCARGOS) spobras\$BENEFÍCIOS<-as.double(spobras\$BENEFÍCIOS)</pre>

SEPARANDO DATASET POR TIPO DE VÍNCULO EMPREGATÍCIO

```
cncm<-spobras %>% select(NOME,CARGO,VENCIMENTOS,`DETALHE VÍNCULO`) %>%
  filter(`DETALHE VÍNCULO` == "Concursado NÃO ocupante de cargo em comissão") ## 1371 OBSERVAÇÕES

ccc<-spobras %>% select(NOME,CARGO,VENCIMENTOS,`DETALHE VÍNCULO`) %>%
  filter(`DETALHE VÍNCULO` == "Concursado ocupante de cargo em comissão") ## 43 OBSERVAÇÕES

ncnc<-spobras %>% select(NOME,CARGO,VENCIMENTOS,`DETALHE VÍNCULO`) %>%
  filter(`DETALHE VÍNCULO` == "Não concursado NÃO ocupante de cargo em comissão") # 504 OBSERVAÇÕES

nccc<-spobras %>% select(NOME,CARGO,VENCIMENTOS,`DETALHE VÍNCULO`) %>%
  filter(`DETALHE VÍNCULO` == "Não concursado ocupante de cargo em comissão") ## 444 OBSERVAÇÕES
```

DISTRIBUIÇÃO DE SERVIDORES POR VÍNCULOS E SALÁRIOS

```
# PROPORCAO DE SERVIDORES POR VINCULO
proporcao_servidores<-data.frame(Vinculo=c("Concursado Não Ocupante CM","Concursado Ocupante CM","Não Concursado Não Ocupante CM","Não Concursado Ocupante CM", "Total"),
  Qtd_servidores=rbind(nrow(cncm),nrow(ccc),nrow(ncnc),nrow(nccc),nrow(spobras)),
  Proporcao=rbind((nrow(cncm)/nrow(spobras)*100),(nrow(ccc)/nrow(spobras)*100),(nrow(ncnc)/nrow(spobras)*100), (nrow(nccc)/nrow(spobras)*100),(nrow(spobras)/nrow(spobras)*100)))

### PROPORCAO SALARIO BRUTO POR VINCULO
proporcao_salarial_X_vinculo<-data.frame(Vinculo=c("Concursado Não Ocupante CM","Concursado Ocupante CM","Não Concursado Não Ocupante CM","Não Concursado Ocupante CM", "Total"),
  Salario_Bruto_total=rbind(sum(cncm$VENCIMENTOS),sum(ccc$VENCIMENTOS),sum(ncnc$VENCIMENTOS),sum(nccc$VENCIMENTOS), sum(spobras$VENCIMENTOS)),
  Proporcao_salario_Bruto=rbind(sum(cncm$VENCIMENTOS)/sum(spobras$VENCIMENTOS)*100, sum(ccc$VENCIMENTOS)/sum(spobras$VENCIMENTOS)*100,sum(ncnc$VENCIMENTOS)/sum(spobras$VENCIMENTOS)*100, sum(nccc$VENCIMENTOS)/sum(spobras$VENCIMENTOS)*100,sum(spobras$VENCIMENTOS)/sum(spobras$VENCIMENTOS)*100) )
```

TRANSFORMANDO EM TIBBLES

```
proporcao_servidores<-as_tibble(proporcao_servidores)
proporcao_salarial_X_vinculo<-as_tibble(proporcao_salarial_X_vinculo)
```

TABELA E GRAFICO: PROPORÇÃO SERVIDORES POR VINCULO

```
proporcao_servidores
```

Vinculo <chr>	Qtd_servidores <int>	Proporcao <dbl>
Concursado Não Ocupante CM	1371	58.044030
Concursado Ocupante CM	43	1.820491
Não Concursado Não Ocupante CM	444	18.797629
Não Concursado Ocupante CM	504	21.337849
Total	2362	100.000000

5 rows

```
proporcao_servidores %>%
  ggplot(mapping = aes(x=Proporcao,y=Vinculo,fill=Vinculo))+
  geom_col(show.legend = FALSE)+
  ggtitle(label = "Proporção Servidores por Vínculo Empregatício")+
  xlab(label = "Porcentagem (%)")+
  ylab(label = "Vínculo Empregatício")
```

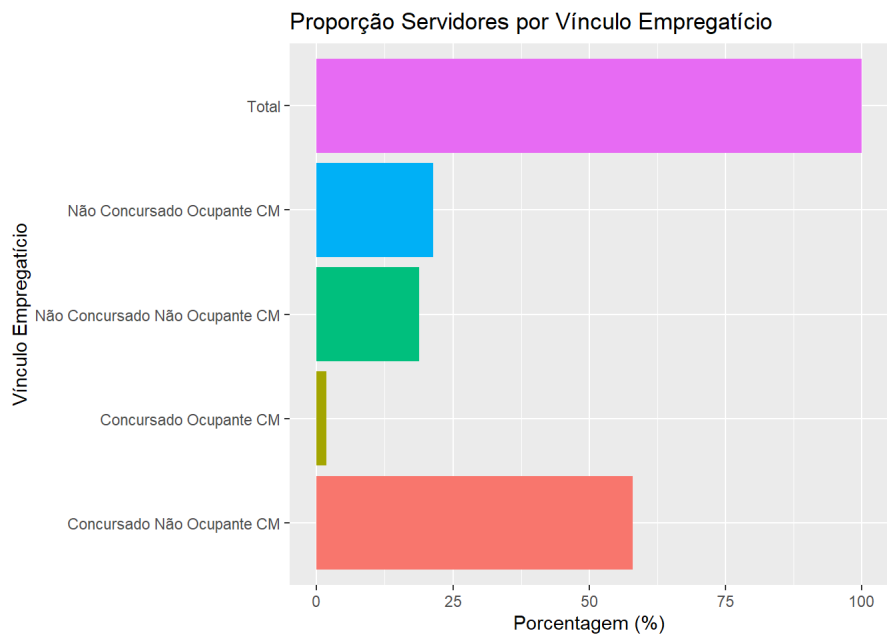
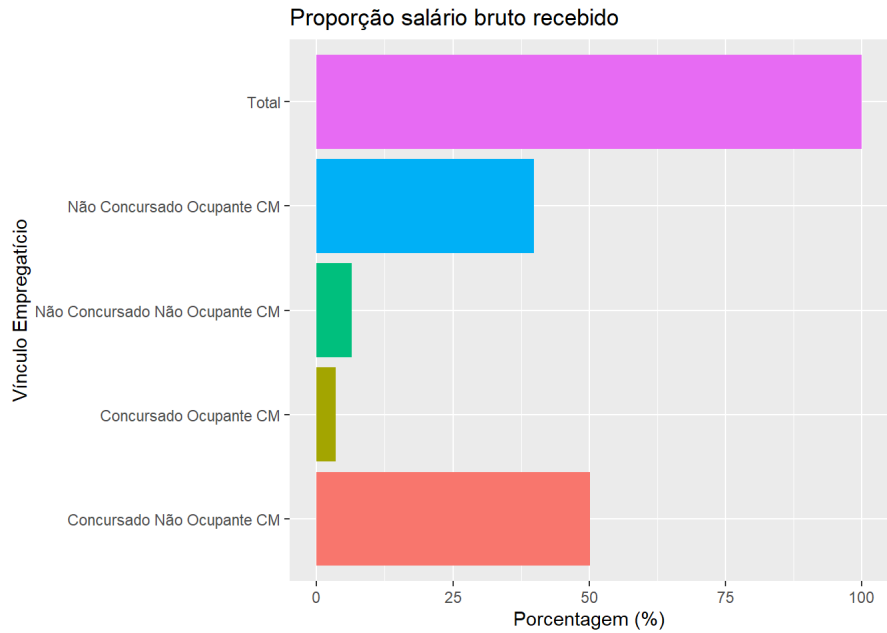


TABELA E GRAFICO: PROPORÇÃO SERVIDORES SALARIO BRUTO POR VINCULO

proporcao_salarial_X_vinculo		
Vínculo <chr>	Salario_Bruto_total <dbl>	Proporcao_salario_Bruto <dbl>
Concursado Não Ocupante CM	11042321	50.175480
Concursado Ocupante CM	768121	3.490285
Não Concursado Não Ocupante CM	1416981	6.438654
Não Concursado Ocupante CM	8779982	39.895581
Total	22007405	100.000000
5 rows		
<pre>proporcao_salarial_X_vinculo %>% ggplot(mapping = aes(x=Proporcao_salario_Bruto,y=Vinculo,fill=Vinculo))+ geom_col(show.legend = FALSE)+ ggtitle(label = "Proporção salário bruto recebido")+ xlab(label = "Porcentagem (%)")+ ylab(label = "Vínculo Empregatício")</pre>		



TRANSFORMANDO DATA DE ADMISSÃO EM DATA RECONHECIDA PELO R

```
spobras$ADMISSÃO<-dmy(spobras$ADMISSÃO)
spobras %>% select(ADMISSÃO)
```

	ADMISSÃO <date>
	1976-08-19
	1979-10-12
	1980-04-10
	1980-07-14
	1981-08-03
	1981-09-17
	1981-12-01
	1981-12-28
	1983-11-01
	1987-11-18

1-10 of 2,362 rows

Previous 1 2 3 4 5 6 ... 237 Next

CRIANDO UMA VARIÁVEL COM OS ANOS DE TRABALHO DE CADA SERVIDOR.

```
spobras<- spobras %>% mutate(ANOS_DE_TRABALHO=(today()-ADMISSÃO)/dyears(1))
spobras$ANOS_DE_TRABALHO<-as.integer(spobras$ANOS_DE_TRABALHO)
spobras[,c(4,7,15)]
```

NOME <chr>	ADMISSÃO <date>	ANOS_DE_TRABALHO <int>
WALTER MARTINS DE SOUZA	1976-08-19	44
LUIZ CARLOS DE SOUZA BARBOZA	1979-10-12	41
VALDELI DA SILVA	1980-04-10	40
MARCIA APARECIDA TEIXEIRA DE CAMPOS CARVALHO	1980-07-14	40
MARCOS SILVA ROCHA	1981-08-03	39
JOAO PAULO RESPLANDES DE OLIVEIRA	1981-09-17	39
ELIANA MARIANO FRANCO	1981-12-01	38
SILVANA ISGROI CARVALHO	1981-12-28	38
MARCIA HELENA NUCCI EIRAS GARCIA	1983-11-01	37
ALFREDO LENZ	1987-11-18	32

1-10 of 2,362 rows

Previous 1 2 3 4 5 6 ... 237 Next

MEDIA SALARIAL e DESVIO PADRAO POR CARGO OCUPADO

```
spobras %>% group_by(CARGO) %>% summarize(media_salarial=mean(VENCIMENTOS,na.rm = TRUE),
                                           desvio_padrao=sd(VENCIMENTOS,na.rm = TRUE),Qtd=n()) %>% arrange(desc(media_salarial))
```

```
## `summarise()`` ungrouping output (override with `.groups` argument)
```

CARGO <chr>	media_salarial <dbl>	desvio_padrao <dbl>	... <int>
DIRETOR DE OBRAS	30062.5000	23125.00000	4
CHEFE DE GABINETE	24459.5150	10622.07474	12
GESTOR SENIOR	22533.7763	6541.20392	24
DIRETOR PRESIDENTE	22483.9038	12925.31866	13
DIRETORA DE REPRESENTAÇÃO DOS EMPREGADOS	22354.1667	13351.22498	12
DIRETOR ADMINISTRATIVO E FINANCEIRO	22327.0369	15547.67078	13
DIRETOR DE PROJETOS	21292.7400	12262.06693	5
COORDENADOR DE NUCLEO	19178.5074	6949.76485	162
COORDENADOR DE AREA PLENO	19031.6012	4346.94135	24
DIRETOR TECNICO	18500.0000	0.00000	8
1-10 of 44 rows	Previous	1	2 3 4 5 Next

MAIOR SALARIO MENSAL RECEBIDO ENTRE TODOS OS SERVIDORES

```
spobras %>% select(NOME,ANOS_DE_TRABALHO,VENCIMENTOS) %>% arrange(desc(VENCIMENTOS))
```

NOME <chr>	ANOS_DE_TRABALHO <int>	VENCIMENTOS <dbl>
MAX YAMATOGUE FAGUNDES	3	79027.51
PAULO SANTORO DE MATTOS ALMEIDA	3	70633.67
SELMA BARROS DOS SANTOS DIAS	3	64750.00
GIOVANI OLIVEIRA DA COSTA	2	64750.00
MAURICIO BRUN BUCKER	2	60801.71
VIVIAN PATRICIA BANDEIRA NOGUEIRA PINTO	2	41530.80
MARINA A KYRIOPOULOS	3	40585.80
HILDA MITIKO IUAMOTO PACHECO	3	39682.58
HILDA MITIKO IUAMOTO PACHECO	3	39138.97
NATALIE MATA	4	38567.67
1-10 of 2,362 rows	Previous	1 2 3 4 5 6 ... 237 Next